AIR WAR COLLEGE

AIR UNIVERSITY

INTELLIGENCE, SURVEILLANCE, AND RECONNAISSANCE PROCESSING, EXPLOITATION, AND DISSEMINATION SYSTEM IN SUPPORT OF GLOBAL STRIKE IN 2035

by

Erik C. Bowman, Lt Col, USAF

A Research Report Submitted to the Faculty

In Partial Fulfillment of the Graduation Requirements

15 Feb 2012

DISCLAIMER

The views expressed in this academic research paper are those of the author and do not reflect the official policy or position of the US government or the Department of Defense. In accordance with Air Force Instruction 51-303, it is not copyrighted, but is the property of the United States government.

Biography

Lieutenant Colonel Erik C. Bowman, USAF, is currently a student at the Air War College, Maxwell AFB, AL. Prior to this assignment, he commanded two space launch squadrons and served as the 45th Launch Group deputy at Cape Canaveral AFS, FL. He is a career developmental engineer, but has held a variety of positions in fields ranging from ballistic missile and space intelligence, academia (Assistant Professor of Astronautics at the US Air Force Academy), space superiority requirements and program management, as well as space launch operations. He has also served two tours as a UN Weapons inspector in Baghdad, Iraq in 1995 and 1996. Commissioned in 1991, he holds a BS in Astronautical Engineering from the US Air Force Academy, MS in Aeronautical/Astronautical Engineering from Purdue University, and an MS in C4I systems from the Air Force Institute of Technology.

Abstract

"Strike" is not the main challenge facing Global Strike in 2035—it is to globally attribute, target, and assess the strike. The United States cannot afford worldwide sensors to perform this role, but it does not have to. Much of the information required is already being collected—the US just needs to grab it, process, exploit, and disseminate it. This means the US needs the ability to connect to (currently) undiscovered data sources and process that data in currently unanticipated ways to support unanticipated users. One system that is good at processing new information in new ways is the human brain. If the US could design its Processing, Exploitation, and Dissemination (PED) architecture to function similarly to the human brain, the system could support unanticipated needs more easily. The power of this "brain" to support global strike also makes it a center of gravity for the US. Thus, the PED "Brain" must be built with integrated defenses from the start to prevent its compromise or destruction.

If the interface, data, and security standards can be constructed up front, the system can evolve incrementally to fit within whatever budgets are available, and eventually can become self-constructing. The design enables this behavior by segregating all processing and data into distinct nodes which mimic the behavior of neurons in the brain, where the links between nodes may be more important than the information and processing within each node itself. As more nodes are added, the intelligence of the system increases exponentially. With the exception of some protection technologies, the technology exists today. The primary challenges to implementation are non-technical, but they are significant. These challenges must be addressed in order to implement the PED system required to support Global Strike in 2035. Global Strike will be ineffective if it does not know what it should strike—and the PED system will provide that knowledge.

Introduction

Global Strike is intended to deter and strike adversaries worldwide within operationally relevant timeframes. Today, with the exception of nuclear intercontinental ballistic missiles (ICBM), submarine-launched ballistic missiles (SLBM), and some limited cyber options, the United States has little capability for Global Strike without the need for forward basing or significant logistics support. However, by 2035, the ability to strike will not be the limiting factor. Rather, the primary difficulties will be to globally attribute attacks against the United States, identify the enemy's kill chain, provide the information supporting the US kill chain, and assess the 2nd and 3rd order effects of a strike.

The key to effective Global Strike is to get inside of the adversary's Observe, Orient, Decide, and Act (OODA) loop. Once attacked, the US must be able to take the next escalatory step before the adversary can regain the initiative. Unfortunately, globalization will have leveled the playing field. The adversary will probably have some of the same strike technologies that the United States has, including the ability to strike quickly once a decision to do so is made. This means that the OODA loop will either need to collapse to an OODA point, or the US will need to accurately predict adversary actions before they occur—the concept known as Predictive Battlespace Awareness (PBA). Meeting these timelines is beyond human capabilities. As a result, the information processing, exploitation, and dissemination (PED) system must help decision-makers to think faster and predict adversary actions.

Global strike will most likely be employed to address one of three scenarios:

- Within an existing conflict, the US must **strike** a specific known target quickly wherever it is on the globe. One of the most stressing examples is striking a fleeting, mobile target such as a mobile missile launcher.

- The US must respond to an unexpected attack by **disrupting** the enemy's kill chain to prevent a second attack before it occurs. One of the most stressing examples is preventing a second cyber-attack against US electrical or financial infrastructure.
- The US wishes to **deter** adversaries from attacking. For deterrence to be effective, the US must also demonstrate the capability for Global Strike. One of the most stressing examples is deterring non-state actors such as terrorist cells.

In the first scenario, the goal of the system is to provide all of the information required to support the US' kill chain for the chosen target, including providing information on external defenses. In the second and third scenarios, the system must attribute the attack to a specific enemy. But that is not all. In the second scenario, the system must also identify the kill chain of the attacking weapon to enable Global Strike to disrupt/deny future attacks. In contrast, for the third scenario the system must provide information on enemy links, nodes, and capabilities to aid in selecting a coercive target.

Global attribution, targeting, and assessment also implies the need for a worldwide suite of sensors to watch everything, all the time. Skeptics might rightfully point out that the United States cannot afford to build a worldwide sensor network to observe everything, all of the time, particularly in an era of declining defense budgets. Fortunately, the US does not need to bankrupt itself paying for sensors all over the world – in many cases, the data already exists.[1] One must to tap into a variety of currently inaccessible information sources[2] that collect the information we need, and fuse that information into a coherent picture of what is going on around the world. Although this data is accessible to ***someone***, it is not necessarily accessible to the US military. Consequently, the bulk of the PED effort will be to gain access to this data.

Ideally, this would be accomplished by entering into service level agreements with owners of the data to gain access and translate it into a form that a data fusion/mining engine can understand, but it could also be accomplished by more clandestine means. The data fusion/mining piece is the easy part, relatively speaking. Once that is accomplished, the US needs to connect to those information sources, process data into an understandable form, exploit the information, and disseminate it to intelligence consumers, including the Global Strike community.

Today, the process for adding new sources to the system is cumbersome, slow, and dysfunctional.[3] In 2012, the United States has no way of accurately predicting exactly which information sources and algorithms will be required in 2035. The US also cannot predict which users might need access to the PED system in 2035.

Therefore the PED system must support the ability to connect unanticipated data sources, process that data in unanticipated ways, and support unanticipated users. Data must be discoverable. This sounds difficult, but it is not as hard as it might first appear. The human brain is a biological system that can support these types of vague requirements. Therefore, a system modeled on the human brain might be successful.

The human brain is designed to incorporate new data sources and ways to process that information to support applications that were never considered—to humans, it is simply called learning. The brain is composed of individual neurons that process and store a small amount of information. By forming connections to other neurons, additional information gets stored and processed in the linkages between neurons. The whole is greater than the sum of its parts. And because much of that information is in the linkages, and linkages are easy to break apart and put back together in different ways, the brain has amazing flexibility in learning new things: dealing with unanticipated data, processing it unanticipated ways to meet unanticipated needs. The brain literally rewires itself to answer a new question.

A PED system which emulates the brain would be extremely powerful. It would be able to identify and connect to any information source it needed to answer a particular question. For a surprise attack, it could go back and look at historical records, and look for common threads that trace back to the actual attacker. It could take sparse data and overlay a variety of different paradigms or thought patterns to process data into a more human-intelligible form. It could "learn" how enemy systems operate and predict 2nd and 3rd order effects of a particular attack. In other words, it would be crucial to Global Strike. It is hard to imagine Global Strike being effective without it.

Saying that, its power will make it a target. If adversaries view this system as a center of gravity, they will attempt to destroy it, deny its use, degrade it, disrupt it, and may even attempt to deceive it. Therefore, the PED system of 2035 must be designed with integrated defenses from the start. The choice of a brain as the analog for the system is a good start. There are numerous examples of how the brain degrades gracefully when damaged. For example, strokes kill large portions of the brain, but stroke victims gradually regain the ability to speak or use both sides of their body over time as the remaining half of the brain rewires itself to replace lost functionality. Blind people find their other senses are enhanced as the brain reuses the optical cortex to enhance auditory and tactile processing. Even before the brain begins to rewire itself, undamaged portions continue to function—there is not a single point of failure.

These types of defenses are not sufficient. Human brains are also subject to deception. Optical illusions and confidence games can trick the brain into seeing or thinking something that is not true. The same could be true for an electronic brain. In addition, an electronic brain could be destroyed non-kinetically through phenomena such as electromagnetic pulse (EMP) or high power microwaves (HPM).

Thus, this paper asserts that the Intelligence, Surveillance, and Reconnaissance (ISR) PED architecture will need to be re-structured like a brain with integrated defenses—and that this is achievable--to support Global Strike in 2035. This paper will first examine in more detail how the human brain processes information and the traits that might be applied to an electronic analog of the brain. It will then describe a concept of operations for how the system will translate the human brain's mode of operation into a man-made architecture that emulates the brain's functionality—a notional system architecture. It will also lay out a notional roadmap to evolve such a system to its desired end state. Finally it will make some recommendations and draw conclusions.

The Human Brain

To actually enable Global Strike to address the three scenarios described in the introduction is an extremely daunting problem. Today's PED system is certainly not up to the challenge. How can the US hope to actually build something omniscient enough to link and fuse all of these data sources to support Global Strike under these scenarios? The answer is simple in theory– the PED system should be constructed to operate more like the human brain, but should operate much more quickly.

The human brain is not a monolithic, homogenous processing center. It is composed of approximately 10 billion nerve cells, also known as neurons[4]. Each neuron (figure 1) is composed of dendrites (receive information), the cell body (stores and processes information), and the axon (delivers information to other neurons). Inside the human brain, dendrites receive electrical inputs (which may number in the thousands) from other neurons. The cell body performs an operation similar to a sum, and when a certain threshold is achieved, it discharges an electrical impulse down its axon to other neurons.[5] This is akin to a computer receiving an input, performing some computation, and sending that information to some other computer to which it is connected.

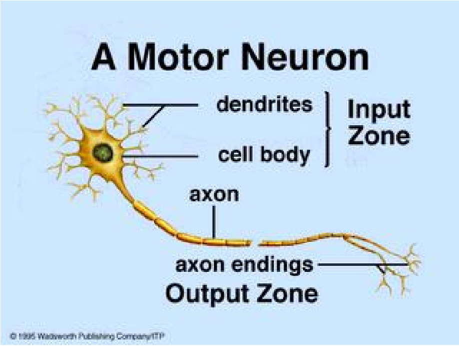


Figure 1: Architecture of Neuron[6]

The links or connections between neurons are called synapses (figure 2). The way brains learn is by varying the intensity of the connections between neurons, adding neurons to the connections and deleting them. The more times a synapse is activated by a neuron, the stronger the connection gets.[7] Learning can also occur if multiple neurons are activated by the same stimulus.[8] In other words, experience is equivalent to multiple activations of the synapse (the more times one encounters something, the more experience one gains). The stronger the connection, the stronger the memory. Fundamentally, the brain stores memories in the synapses or connections between neurons as well as within individual neurons themselves.

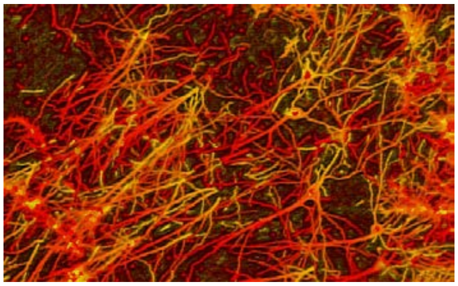


Figure 2: Nerve synapses inside the brain[9]

While each neuron only has a very small processing and memory storage capability, the aggregation of neurons and the (potentially) thousands of connections each neuron makes with other neurons results in a powerful processing capability. Our overall consciousness is an aggregation of all the neurons of the brain communicating. For all intents and purposes, we only think as a single monolithic brain (figure 3), even though our consciousness is the aggregation of billions of individual interconnected neurons.

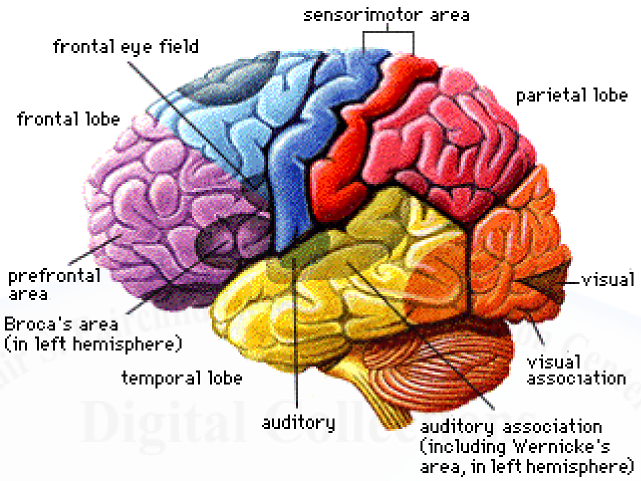


Figure 3: The Human Brain[10]

This structure of the human brain has several useful characteristics that may support the PED system of 2035. These include:

- The distributed nature of the brain means that it can degrade gracefully when damaged. [11] The loss of a few neurons out of a billion does not result in catastrophic loss of brain function—just decreased capacity.
- Healthy portions of the brain can “re-learn” what the damaged portions of the brain did to restore lost functionality.[12]
- It performs massively parallel computations (such as recognizing patterns, faces and voices) easily and efficiently.[13]
- It can reorganize its neurons and connections, and thus learn, from experience.[14]

Obviously, if one can construct a system to operate similarly, Global Strike can gain enormous benefits. Since predicting behaviors of an adversary ultimately revolves around learning new information, integrating it, and processing sparse or noisy information into intelligible patterns,[15] such a system should be ideally suited to supporting attribution, targeting, and assessment, and may even be able to provide some advance warning of attack from non-obvious indicators.

CONOPS

How would a system that duplicates this functionality operate? There are a couple of current examples that have tried to emulate the functionality of the brain. The most well-known is the Internet. The Internet networks millions of computers to one another worldwide through a common protocol known as Transport Control Protocol/Internet Protocol (TCP/IP). Its distributed, packet-switched nature makes it extremely robust against disruption and destruction, just like the human brain. However, while the Internet simulates the physical construction of neurons in the brain,[16] it doesn't emulate the learning process—it doesn't automatically link node to node.

Google is an example that begins to link nodes to emulate the way the brain learns. It employs web crawlers to locate bits of information, which are akin to individual neurons of the brain. It then processes the locations of all of this content to generate an index. The generation of the index is akin to neurons making connections to other neurons.[17] Where the learning analogy applies is within Google's PageRank system. This assigns a numerical weighting to each page based on the number of links to each page, and the relative importance of each page.[18] This is akin to the way the brain learns by weighting connections according to the frequency of firings of a neuron. What it cannot do is remember the context attached to those memories and facts. For example, when a user types the word "bunkers" into Google, the top-ranked links discuss hardened and buried shelters. But there are also links to sand traps on the golf course and links to “All in the Family”[19] episodes on YouTube. Which type of bunker is the user interested in? Context is required to answer that question.

Systems like Facebook and Amazon begin to provide some of that context by harvesting it from their users. To extend the example from above, the user might have posted on Facebook that they received some golf clubs for Christmas the previous year. Then in a subsequent post, they state, “I’ve been having a lot of problems with bunkers lately.” Facebook advertising might pop an ad for a golf course that has no sand traps or an ad for a new sand wedge. It uses the context of prior knowledge of the golf club gift to determine the context of the ambiguous term of “bunker”.

The PED system should leverage the best of all of these systems to duplicate the brain’s behavior. Where possible, it should learn directly from stimuli and make associations based on context. Neural nets are a technology that may accomplish this feat. Neural nets create random connections between facts, but these facts are only “learned” when the random connection is verified by a human or machine trainer.[20] The problem with a neural net is that a lot of training is required before it starts to “think” in a useful fashion.

What the PED system needs is a way to jump start this learning process. To that end, it should leverage the physical architecture of the Internet and create initial links intelligently (rather than randomly) through an expert system such as PageRank. Then it should employ context-based neural nets to assess the quality of the information and refine the links to provide the most relevant information.

This is exactly how the brain learns. The links described above are analogous to synapses. Learning occurs when synapses are activated multiple times, or when multiple synapses are excited at the same time by the same stimulus, strengthening the connections between neurons. This is how humans learn from experience. The stronger the connection, the stronger the memory, or learning that has occurred.

From a functional standpoint, the PED “brain” must be able to both process information and store information. In a human brain, neurons both store and process information, but the PED analog could also operate effectively with the functions separate, as long as storage and processing are tied together in some manner. Of course, the data has to be represented in a form a computer can understand, and output in a form that is as intuitive for humans as possible.

Sadly, there are other non-technical challenges that must be overcome before the US can build a system that emulates the human brain. The first of these are to break down stovepipes between information sources. A brain that could not share information between the cerebrum and cerebellum, or between the left and right halves, is a brain that is crippled in functionality. The same is true for a PED system. Information that is available to one agency is not necessarily shared with another agency for security or bureaucratic reasons.[21] This is enforced either via security compartmentalization or by purposeful incompatibility of computer systems and networks.[22]

The second challenge is organizational conflict of interest -- sensors and associated data processing are often produced by different contractors. Thus, sharing data, as well as the development of the system itself, could result in proprietary information being divulged to business competitors.[23]

The third challenge is the certification and accreditation (C&A) process. The human brain does not need approval from a third party before it incorporates information, but Department of Defense (DoD) computer systems require this approval for each new application and data source connected to DoD networks. Unfortunately, the certification process is highly centralized and time-consuming.[24] Furthermore, if the system truly functions like a brain, it will dynamically connect and disconnect to new data sources and applications autonomously, but the C&A process is not currently structured to handle the network dynamically configuring itself.

Fourth, cyber techniques can be exploited by adversaries to surreptitiously extract, modify, or destroy information from computer systems. In contrast, the human brain can be destroyed, but it is not currently possible for an enemy to eavesdrop on individual neurons or collections of neurons to surreptitiously recover memories. It is also not currently possible to surreptitiously alter or delete individual memories in a person's brain.[25] Therefore, the US will need to expand the system to include not just the traits of the human brains, but also include security. It should be able to detect attempted intrusions into the system, and it should also be able to detect unauthorized attempts to exfiltrate or modify the data by insiders. Additionally, the system should identify the authoritative source for each type of information, and provide mechanisms for that source to periodically verify the accuracy of that data.

Lastly, there is inherent conflict with the US Bill of Rights and subsequent laws that have been passed such as the Wiretap laws and the Foreign Intelligence Surveillance Act (FISA) governing the collection of information on US citizens. For the system to be truly effective, laws will have to be changed to dissolve the legal borders between DoD and domestic intelligence responsibilities. In order to attribute an attack against the US and trace it to the source, both domestic and foreign sources must be accessed and fused. Civil and private information must be linked with military databases. This causes severe conflict with the 4th amendment to the US Constitution against unreasonable search and seizure, and the 5th amendment associated with Due Process. Whatever system is developed to support global strike will need to provide sufficient information to attribute and respond to attacks against the US anywhere on the planet, but must also protect the rights of US citizens in the process. Now that the desired functionality of the PED system has been described, what could this PED brain actually look like?

Design of the PED Brain

Physically, the PED brain can be described as an aggregation of 3 different types of nodes. Each node can be thought of as an individual neuron in the brain (figure 4).

- Finished Data nodes—nodes that store data along with its contextual information and links to other information. These are represented by the yellow circles.
- Processor Nodes – nodes that process raw data from a sensor, finished data nodes, System Control and Data Acquisition (SCADA) systems, human beings, custom applications, or external database or web page, and convert it into data directly usable by the PED system and its end users. These are represented by the green circles.
- Security nodes – nodes that authorize access to data and processing algorithms. This is represented by the black rectangle at the bottom of figure 4.
- Links -- The links between each node (represented by the solid arrows showing the direction of information flow) can be thought of as the synapses in the brain which connect neuron to neuron.
    - Every link between neurons must be authenticated and authorized by the security node to allow data to flow. This is represented by the dotted lines in the diagram.[26]
    - Note that links can be recursive. They can link to themselves or go backwards (finished data feeding into a processor node, as well as sensor data).

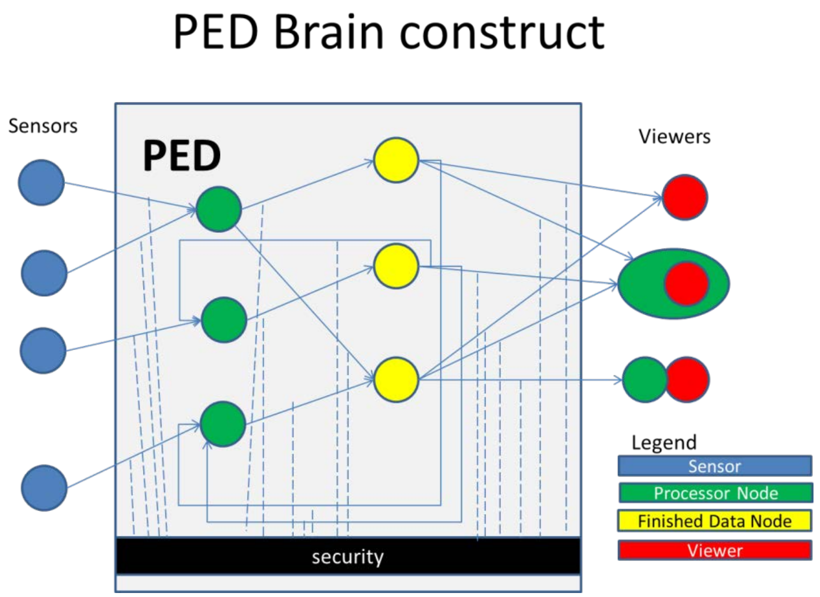


Figure 4: PED Brain construct

Note that everything outside of the gray area in the figure--sensors, external databases, custom applications, human interactions, etc.--is considered outside of the boundary of the PED system. The system must interface with these entities to process their inputs and send its outputs, but they are not part of the PED architecture itself.

The ideal user interface would simply be "viewers" that display the finished data from the PED in a form easily intelligible for a user, and it could be displayed in different ways. Even though the viewers are outside of the bounds of PED, the initial implementation would probably include some generic viewers to provide a baseline tool which users could customize. For example, a person could view positions of aircraft as a text list of latitudes and longitudes, translate that into a street address, display the locations visually on a map or satellite image, or view them as locations on or above a 3-D globe view. If the data does not exist in a form that is directly viewable, the user is always free to process or translate that data into a form that they desire. This type of scenario is depicted as the red dot inside the green oval (outside of the gray PED box). However, if the users discipline themselves into separating the processing from the viewing (as shown with the red and green circles immediately next to one another), an additional benefit accrues. Whatever processing algorithm that was developed by a particular user can be shared with the rest of the community and absorbed into the PED itself. Absorbing additional algorithms, and more and more links, can result in the PED “learning” over time. The more processing nodes, data, and links between them that become incorporated into the PED architecture, the “smarter” it becomes. This allows the PED brain to evolve -- it does not have to be created all at once.

This paper has now described how the proposed architecture can begin to look and function like a brain. How can one address some of the other concerns addressed earlier, such as authentication, authorization, confidentiality, and robustness? To do that requires looking at the security node.

As pointed out above, every node requires authorization from the security node to connect to any other node. The operation of the security node and the other PED nodes operates on the model of third-party introductions.[27] The security node can be thought of as a massive security clearance database similar to the Joint Personnel Adjudications System (JPAS) that authenticates a user. The security node will identify the lowest common clearance among user, user’s computer, PED system nodes, and communication path and generate a token or ticket stamped with the approved classification level. In practice, this would all be enabled via a combination of different types of encryption keys.[28] Once that ticket is issued, the user can proceed to connect to those resources for a specified period of time before having to re-authenticate to the security node.

The key enablers are to tag all data with a classification code, a timestamp, and for every user and computer to have an assigned clearance level.[29] Furthermore, as more and more nodes proliferate, there will need to be some type of index that allows users and other nodes to know where the information resides. However, there are some single points of failure – the security node and the index. If either of those nodes are destroyed, corrupted, or denied, the usefulness of the system quickly approaches zero.

The answer to that challenge is massive replication of the index and limited replication of security nodes (due to the need to centrally adjudicate clearances). The primary impediment to replication of the index is the burden of link maintenance. This link maintenance requirement drives the need for an automated, trusted algorithm similar to the GoogleBot to maintain the links and update the indices. By 2035, processing speed and storage capabilities should make resolving these challenges a non-issue.[30]

However, one issue that will be problematic by 2035 is the verification and authentication process. Currently, authentication is accomplished by public key cryptography. Generally speaking, this encryption is based on the factoring of extremely large numbers,[31] which is a computationally difficult process. The bigger the number (representing the public key), the longer it will take to reverse engineer the factors (representing the private key), increasing the security of the encryption.[32] It requires a succession of guesses to break this encryption with brute force, and at current computer speeds, this can take a very long time (which is why these techniques are considered secure).

The problem is that new technologies such as quantum computing and DNA computing are ideally suited to solving potentially infinite numbers of factorization problems simultaneously.[33] This means that deriving the (currently) secure encryption keys (i.e. breaking the code) may be accomplished exponentially faster. There are also some newer technologies such as quantum encryption that may counter quantum computation from a confidentiality standpoint. However, quantum encryption cannot solve the authentication problem. Therefore, the PED system may need to identify new methods of public key encryption that do not depend on the factoring of large numbers, but instead revolve around recursive computations such as those employed by fractals and chaotic systems – quantum and DNA computing are no better at solving these types of problems than general purpose computers.[34]

If all of these problems can be addressed, the inherent structure of the brain will not only solve the confidentiality and authentication problems – it will also solve the robustness problem. The proliferation of the index and the distribution of data across multiple nodes will allow the PED to automatically reroute around damaged or denied areas. Obviously the data in the denied areas will be inaccessible, but the damage will be localized. In contrast, today's PED systems have critical links and node. These are information bottlenecks that, if denied or destroyed, would shut down all data flow, effectively blinding the US. By structuring the PED system like a brain, not only will it be able to "learn" over time, it can move "learned" functions to undamaged portions and gradually recover functionality, much as a stroke victim does.

Unfortunately, there is one attack that can overcome even this design – electromagnetic pulse (EMP). Since this system is envisioned as operating on electronic computers, running on electricity, with electrical or RF data connections, all of these underlying hardware mechanisms are vulnerable to electromagnetic pulse attack.[35] This has the potential to destroy all nodes simultaneously. This will be a problem even if key nodes are in hardened and deeply buried facilities. Even electromagnetic shielding will have no impact unless all of the connections are severed prior to the pulse—and a brain with no connections between nodes is no longer a brain. The brain would be dead, permanently. Even if backups and spare parts existed, they would probably be destroyed at the same time as the brain itself unless they were shielded.

Faraday cages around nodes could potentially mitigate these problems, but some forms of EMP that also produce heavy doses of x-rays and gamma rays would bypass Faraday cages. The only defense would be to add a layer of lead shielding in parallel to the Faraday cages and to add fast response surge suppressors to power and network communication lines.[36] Computers redesigned to use alternative technologies such as fiber optics, optical switches, and optical computing can also help resolve this problem. The reason is that magnetic fields induce movement of electrons, which is the mode of destruction for EMP. Any device that carries current is vulnerable to EMP-induced current surges. The induced current from EMP's extremely strong, rapidly changing electromagnetic fields is enormous, literally burning up electronic components designed to handle much lower currents. In contrast, electromagnetic fields may bend light, but changing magnetic fields won't generate high doses of photons that would saturate the computers. Therefore, if the US wishes to retain any level of its electronic data in the event of EMP, this paper recommends further development of optical switching and optical transistor technologies.

Development Roadmap

Building such a system appears to be a huge acquisition challenge, but really it is not. The reason is it can be built incrementally, node-by-node, and gradually grow to increase functionality. The scenarios described in the introduction encompass a variety of "threads". The information and processing required to identify a target is a thread. The information required to support a specific cyber-attack is another thread. Capability can be constructed one thread at a time. One can start by building a "thin" thread (e.g. one leveraging only one data source), and gradually add more data sources and processing to make the threat "thicker". Or one could build more threads that leverage the same sources to build a PED "web" of information. The bottom line is that the system can be built in small increments, thread by thread, source by source, algorithm by algorithm, to provide useful information to the user while the PED "web" as a whole is still under construction.[37]

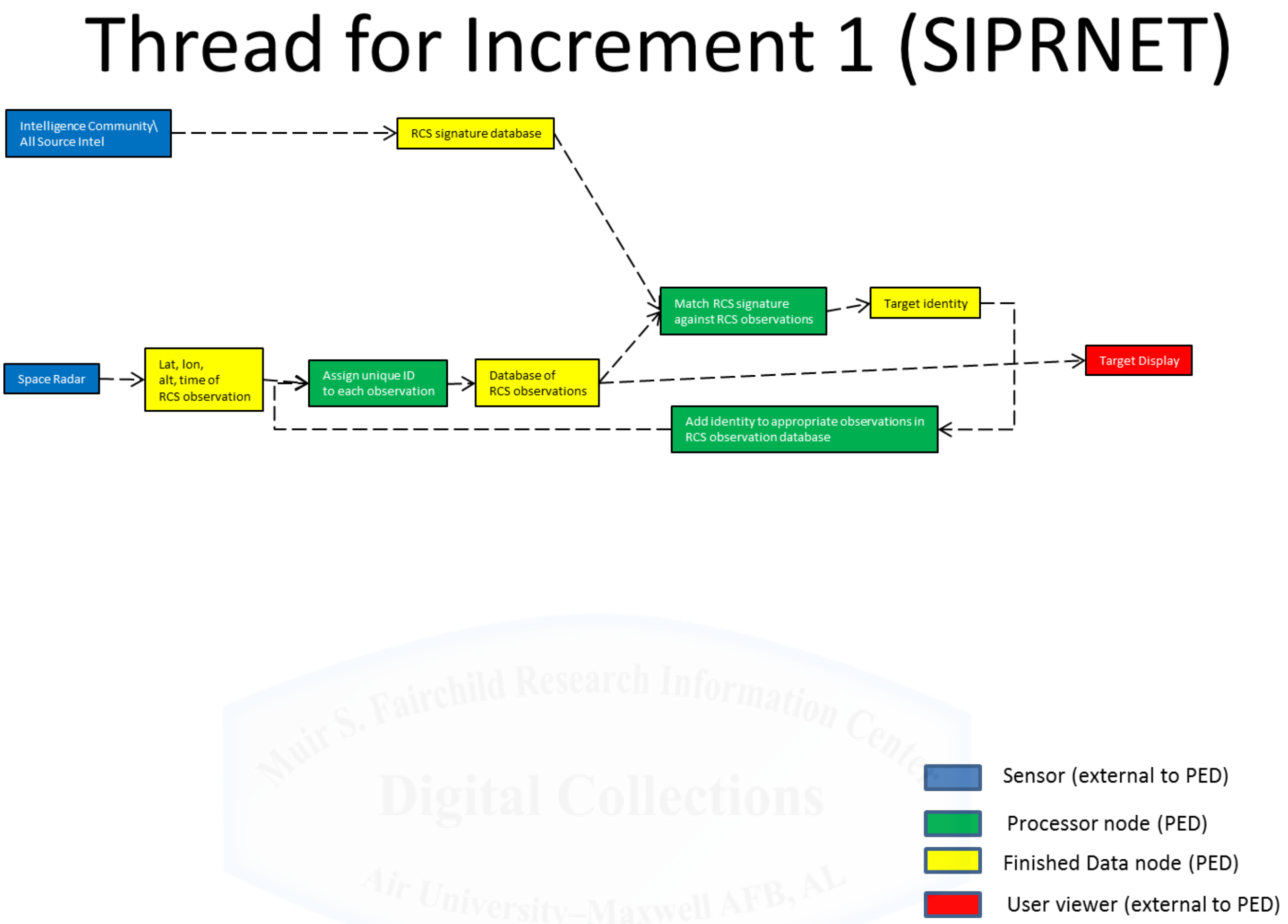
The most critical aspect to building this is defining the standards with respect to how data is represented and how the nodes interact. The PED does not need to know what any of the nodes look like ahead of time, as long as they all work on the same interface. The analogy is the electrical grid. The electrical grid operates at approximately 110V at 60 Hz frequency using alternating current. In addition, wall plugs that deliver this power consist of 2 flat prongs (with one bigger than the other for polarization) and a grounding plug. These are standards and standard interfaces. Because those standards have been defined, an inventor or user can plug any device into the grid and expect it to work, even though these standards were defined over 100 years ago.

The PED system can work in the same manner—define a set of standards and let all of the sensors and user viewers (and new processing algorithms and databases) "plug-in" to the PED using those standards. The acquisition community can start building the network one user and sensor and algorithm and database at a time. To that end, this paper proposes a notional evolutionary schedule.

Increment 1 – Manually Construct the Links

For the first increment, start with a handful of data sources that need to be fused—whatever is required to develop the first thread that is useful to the user community. Identify the classification of each type of data that needs to be produced. Define the data and interface standards for the entire architecture, even those that will not be needed until later increments (e.g. to handle multiple classification levels and compartments across multiple networks).[38]

Appending this type of standard later would cause a nightmare of backwards-compatibility issues that will be enormously expensive to resolve and maintain. Then build a processor node to translate the current raw data into a finished form tagged with the appropriate classification for each piece of data processed. Simultaneously, build a finished data node with a database management system that delivers the maximum possible subset of data a user requests that they are allowed to see for their desired classification level. Also, build a security node at the same time which has the security clearances of a subset of individuals and computer systems that will produce an encryption-protected/authenticated token that specifies the session classification level. Then build a viewer which can automatically request an access token as well as make requests for data from a source. Finally, manually construct an index on each node that points to the pieces of data on the other nodes within the increment 1 system. When this is complete, the user should be able to process at least a single thread of information (for example, identification of a single class of mobile targets using radar cross section (RCS) signatures). One possible implementation of this 1st thread, notionally to identify an unknown object using radar returns on the Secure Internet Protocol Network (SIPRNET),[39] is shown in Figure 5:



Thread for Increment 1 (SIPRNET)

Intelligence Community\ All Source Intel
RCS signature database
Match RCS signature against RCS observations
Target identity
Space Radar
Lat, lon, alt, time of RCS observation
Assign unique ID to each observation
Database of RCS observations
Target Display
Add identity to appropriate observations in RCS observation database

🟦 Sensor (external to PED)

🟩 Processor node (PED)

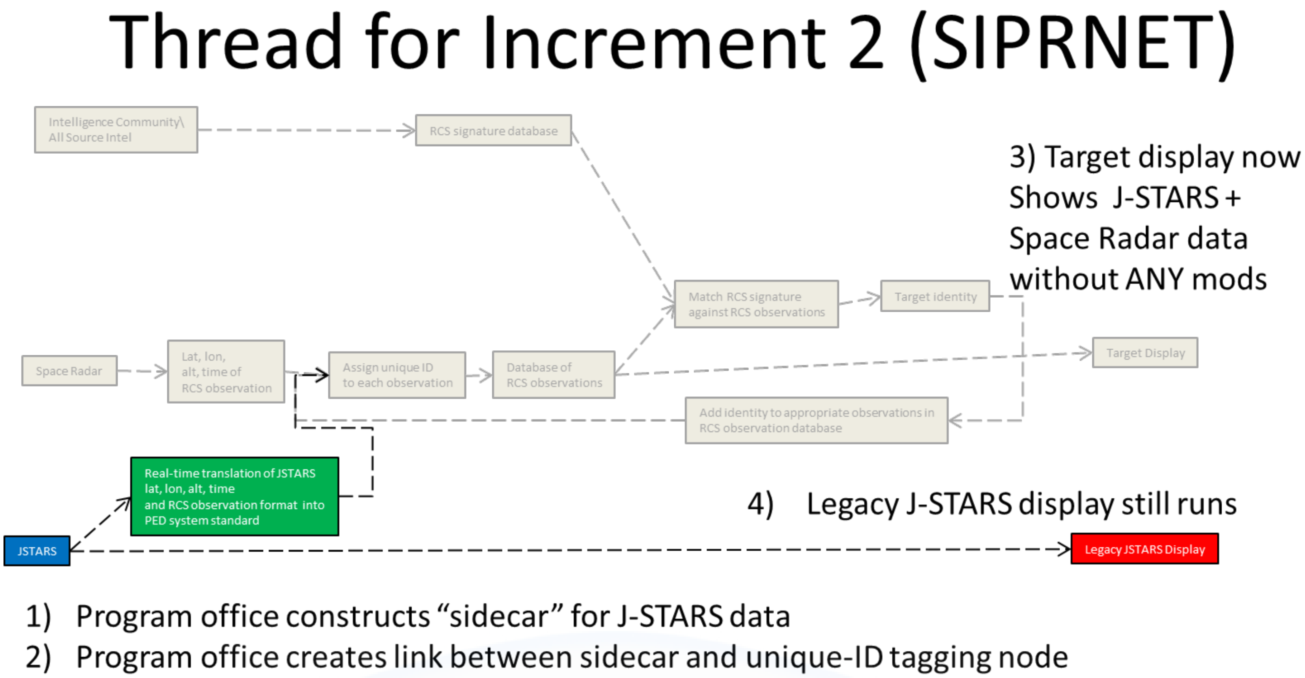
🟨 Finished Data node (PED)

🟥 User viewer (external to PED)

Figure 5: Notional Increment 1 implementation

Increment 2: Integrate legacy systems

The main addition in increment 2 would be to build a translator or sidecar[40] for one or two legacy systems, such as integrating Joint Surveillance and Target Attack Radar System (JSTARS) data. The translator or sidecar is a type of processor node that translates the legacy file format into whatever new data standard was defined in Increment 1.



Thread for Increment 2 (SIPRNET)

Intelligence Community\ All Source Intel

RCS signature database

Match RCS signature against RCS observations

Target identity

Space Radar

Lat, lon, alt, time of RCS observation

Assign unique ID to each observation

Database of RCS observations

Target Display

Add identity to appropriate observations in RCS observation database

Real-time translation of JSTARS lat, lon, alt, time and RCS observation format into PED system standard

JSTARS

Legacy JSTARS Display

3) Target display now Shows J-STARS + Space Radar data without ANY mods

4) Legacy J-STARS display still runs

1) Program office constructs "sidecar" for J-STARS data

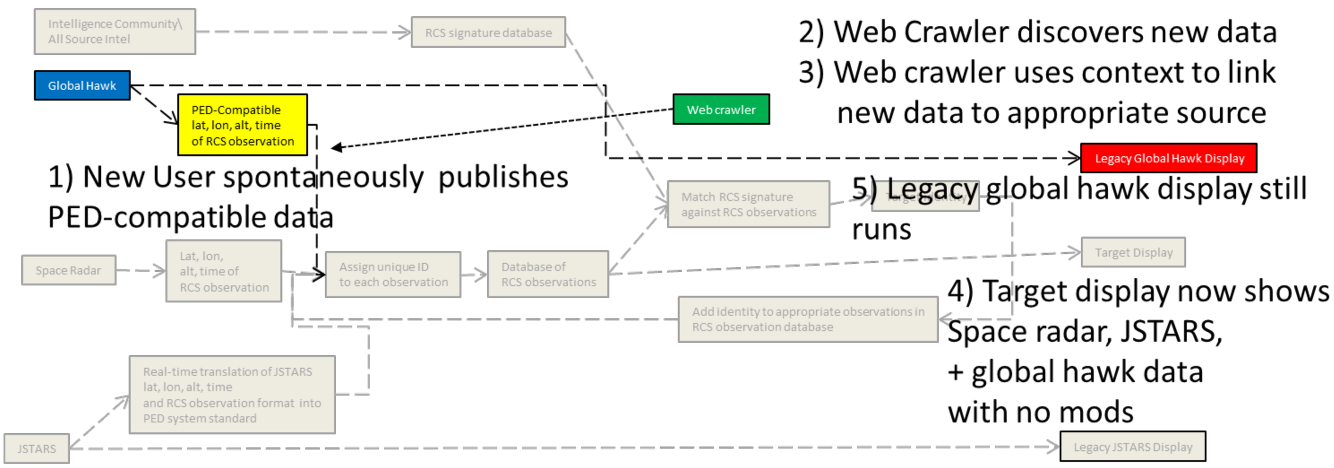
2) Program office creates link between sidecar and unique-ID tagging node

Figure 6: Notional Increment 2 system

Increment 3 – automated maintenance

The primary addition for Increment 3 (figure 7) would be to introduce automated tools for maintaining the index, such as web crawlers[41] and/or neural nets. It may link not just data sources, but start to link processor node to processor node to generate entirely new capabilities. Initially the neural net will require a great deal of human interaction to "train" it into making the right decisions on what information further enhances analysis, Eventually, there may be a "trainer" neural net which will be able to replace direct human judgment to make the net learn faster. But just to emphasize, the neural net for this increment is a nice to have – it is not absolutely essential to have it before moving on to increment 4.

Thread for Increment 3 (SIPRNET)


Intelligence Community\ All Source Intel
RCS signature database
Global Hawk
PED-Compatible lat, lon, alt, time of RCS observation
Web crawler
Legacy Global Hawk Display
Match RCS signature against RCS observations
Space Radar
Lat, lon, alt, time of RCS observation
Assign unique ID to each observation
Database of RCS observations
Target Display
Add identity to appropriate observations in RCS observation database
Real-time translation of JSTARS lat, lon, alt, time and RCS observation format into PED system standard
JSTARS
Legacy JSTARS Display

2) Web Crawler discovers new data
3) Web crawler uses context to link new data to appropriate source

1) New User spontaneously publishes PED-compatible data

5) Legacy global hawk display still runs

4) Target display now shows Space radar, JSTARS, + global hawk data with no mods

Figure 7: Increment 3

Increment 4: Repeat Increments 1-3 on second network

Up to this point, one would construct the PED on a single network such as SIPRNET (Secure internet protocol Router network). However, once the acquisition community has demonstrated its operation, it might repeat the 1st three increments on the Joint Worldwide Intelligence Communication System (JWICS). (Figure 8)

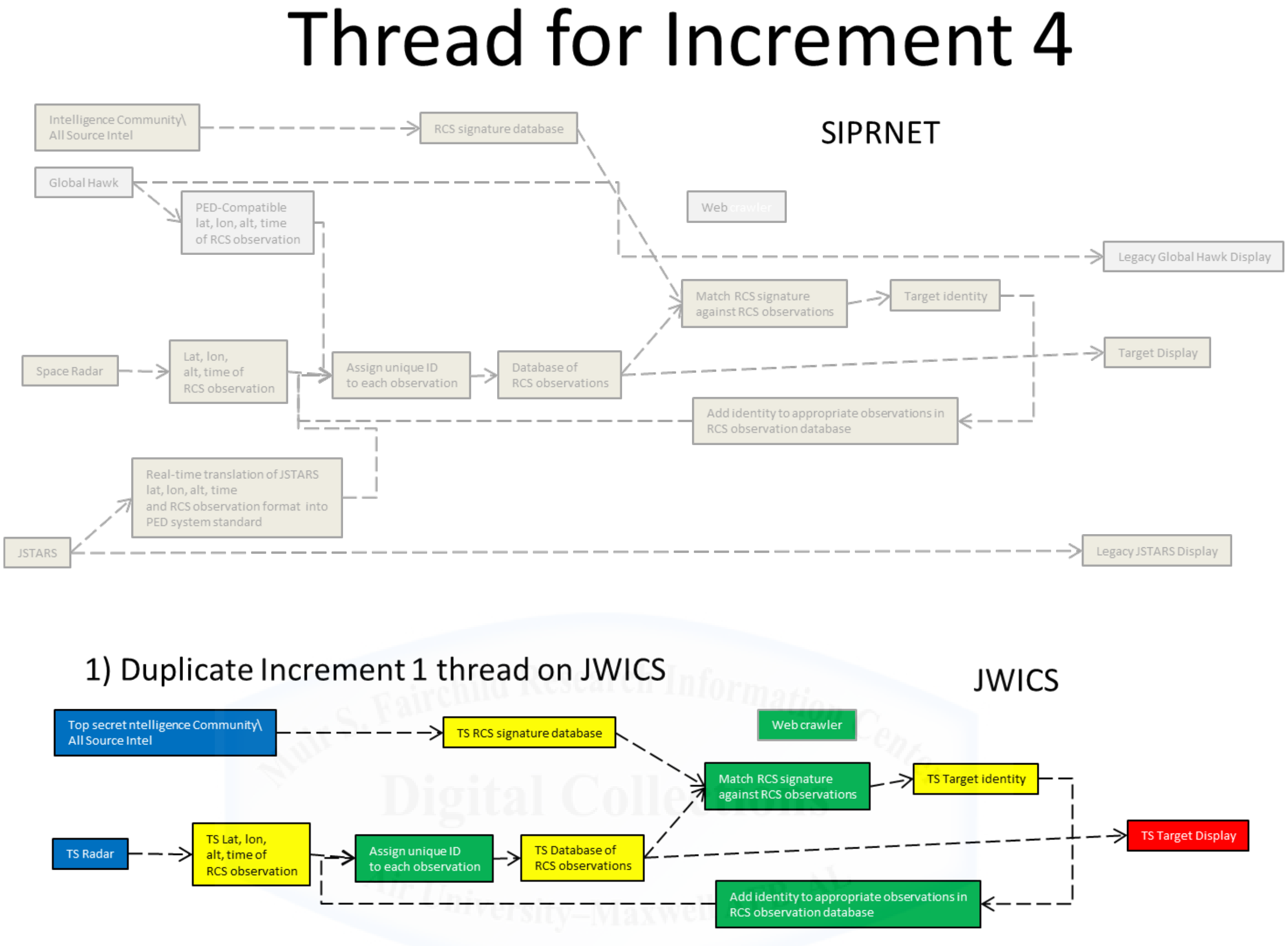
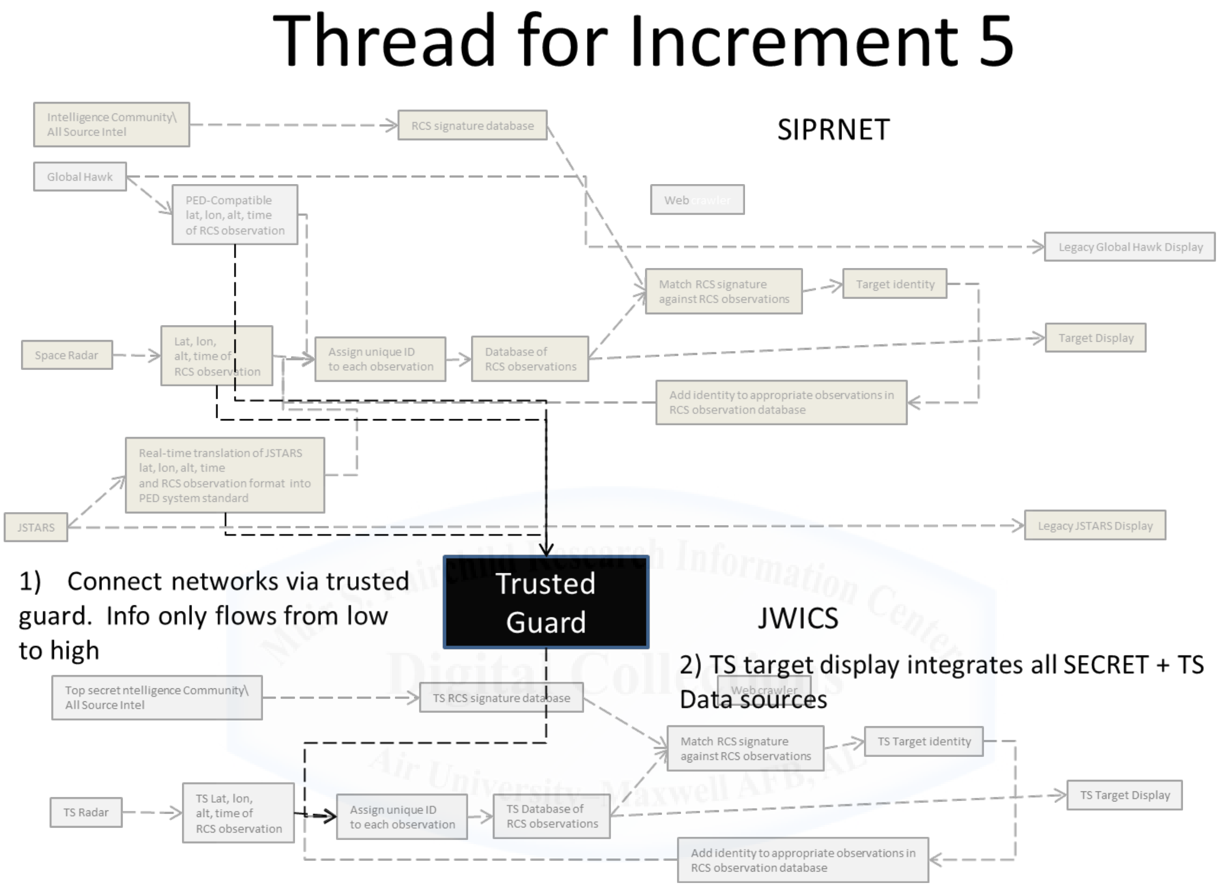

Thread for Increment 4
SIPRNET
Intelligence Community\ All Source Intel
RCS signature database
Global Hawk
PED-Compatible lat, lon, alt, time of RCS observation
Web crawler
Legacy Global Hawk Display
Match RCS signature against RCS observations
Target identity
Space Radar
Lat, lon, alt, time of RCS observation
Assign unique ID to each observation
Database of RCS observations
Target Display
Add identity to appropriate observations in RCS observation database
Real-time translation of JSTARS lat, lon, alt, time and RCS observation format into PED system standard
JSTARS
Legacy JSTARS Display
1) Duplicate Increment 1 thread on JWICS
JWICS
Top secret ntelligence Community\ All Source Intel
TS RCS signature database
Web crawler
Match RCS signature against RCS observations
TS Target identity
TS Radar
TS Lat, lon, alt, time of RCS observation
Assign unique ID to each observation
TS Database of RCS observations
TS Target Display
Add identity to appropriate observations in RCS observation database

Figure 8: Increment 4

Increment 5: Cross-network connectivity

Once the PED is using identical standards and operating on two networks at different classification levels, install a trusted guard between networks that allows a user on JWICS to access SIPR data, while preventing the opposite flow.[42] (Figure 9) This will take some time to get buy-in from all relevant agencies that there will be no Top Secret or compartmented data leakage from JWICS onto SIPRNET—it is potentially the longest pole in the tent. This will also require the same security node to be accessible from both networks to allow cross-network functionality. But if the multilevel security model was successfully demonstrated on SIPRNET during increment 1, and showed no signs of leakage through increment 4, getting approval to cross networks may not be as difficult.



Thread for Increment 5

1) Connect networks via trusted guard. Info only flows from low to high

2) TS target display integrates all SECRET + TS Data sources

SIPRNET
Intelligence Community\ All Source Intel
RCS signature database
Global Hawk
PED-Compatible lat, lon, alt, time of RCS observation
Web crawler
Legacy Global Hawk Display
Match RCS signature against RCS observations
Target identity
Space Radar
Lat, lon, alt, time of RCS observation
Assign unique ID to each observation
Database of RCS observations
Target Display
Add identity to appropriate observations in RCS observation database
Real-time translation of JSTARS lat, lon, alt, time and RCS observation format into PED system standard
JSTARS
Legacy JSTARS Display
Trusted Guard
JWICS
Top secret ntelligence Community\ All Source Intel
TS RCS signature database
Web crawler
Match RCS signature against RCS observations
TS Target identity
TS Radar
TS Lat, lon, alt, time of RCS observation
Assign unique ID to each observation
TS Database of RCS observations
TS Target Display
Add identity to appropriate observations in RCS observation database

Figure 9: Increment 5

Increment 6: Critical Mass

Critical mass occurs when a new capability no longer requires the development of new thread, but only requires a single node. For example, all that is required to add a new sensor is simply to publish data to a particular finished data node because all of the associated processing and dissemination was addressed in earlier increments. Development can now accelerate rapidly.

Increment 7: Self-construction

Eventually, users will begin developing their own processing data or sensors using their own resources to meet the standard. They then just plug it in. When they start structuring their own data and applications to function as data and processor nodes, they can be absorbed into the PED architecture automatically. At this point the PED becomes self-sustaining. Similar self-constructing networks exist today. For example, smartphone users create and post smartphone applications to allow others to use them. Once the system is self-constructing, the program office's role will be reduced to maintaining the standards and underlying infrastructure.

The key point to make for all of these increments is that they do not necessarily need to be done in this order, with the possible exception of Increment 1. This is just one way to approach the problem.

Recommendations

This is all in the realm of the possible. A working prototype demonstrating a 2 node operation of this type of architecture through Increment 2 was constructed in 2005 using only 6 man-months of effort.[43] The primary challenges are cultural, organizational, bureaucratic, and legal. The primary technological challenges are not in implementation, but in protection – challenges which are shared by all information systems today, not just this PED system.

The first recommendation is to eliminate "Need-to-Know" as a general rule, in favor of pre-adjudicated access control for specific types of information when absolutely necessary. This will require discipline to keep data classified at the lowest possible level to allow access by the unanticipated user. Wikileaks is often stated as the argument against this approach. However, this paper argues that although information sharing enabled the leak, it was not the root cause. The root cause was a trusted individual violating that trust. The US has had leaks of highly sensitive information even when information was not broadly shared (e.g. Aldrich Ames, et. al) that one might contend was far more damaging to national security than Wikileaks. While no system can defend completely against purposeful leaks by trusted insiders, the security protocols described here[44] should allow for better detection and prevention of security breaches, even by trusted insiders.

The second recommendation is to develop Increment 1 in-house, without using contractors. The reason is to prevent organizational conflict of interest amongst contractors, while simultaneously allowing the government to iterate on standards before enshrining them in stone (e.g. in a Request for Proposal).

Third, as recommended in the development outline, start small and evolve. This approach lets the PED system employ cost as an independent variable. By breaking down each capability into small, separately costed chunks that can be built in any order, the system has a greater chance of political survival in the forthcoming austere budget environment because some capability is still being produced despite cuts.

Fourth, use translators to integrate legacy systems initially, but put requirements on the systems to directly interface with the PED as part of their sustainment contracts or next increment of capability. Translators work, but there is overhead that will slow the system down.

Fifth, there should be a category of the certification and accreditation process known as “Authority to Develop”. The only reasonable way to integrate systems on the existing networks is to actually integrate and test them on these networks. Trying to build integration tools for real-world systems on a stand-alone system that only emulates the real world simply will not work.[45] It must be developed on the networks themselves, with suitable safeguards negotiated with the

C&A community to ensure that the entire network is not damaged or compromised during development.

Sixth, the US military needs to change the paradigm from expecting the acquisition community to develop all new capabilities to allowing users to develop and share their own fusion tools if they are so inclined and have the skills to do so. The users know their own needs far more than the acquirers, and the information realm is unique (compared to manufacturing) in that skilled individuals can build useful warfighter capabilities. If the US can do so, it can dramatically lower the cost and increase the speed of true integrated common operational pictures.

Seventh, develop new replacement techniques to the current public key cryptographic authentication techniques to combat future threats posed by quantum and DNA computing. This is a task ideally suited to DoD and National Labs, and for innovative contracts run by such organizations as the Defense Advanced Projects Research Agency (DARPA).

Eighth, develop new technologies and/or shielding such as optical transistors to protect our computer networks from EMP. Again, this is a task for DoD and National Labs or DARPA.

Conclusions

Global Strike cannot succeed without a PED system which can attribute attacks, identify the targets, support targeting, monitor engagements globally, and assess those attacks globally. The speed of enemy actions in 2035 will prevent the US from responding effectively unless many responses are executed automatically, shrinking the OODA loop to a point, or even predicting attacks before they happen. Only a PED system structured like a human brain has a hope of accomplishing these tasks.

Building a PED system that functions and learns in a similar fashion to the human brain to support unanticipated users, data sources, and processing algorithms is technically doable today, and can be implemented for reasonable cost if performed in small increments. Furthermore, leveraging the government to integrate existing systems and releasing contractors to only develop new capabilities may be the best organizational and acquisition approach.

Because of the PED's importance, the US must protect it from attack. Without it, Global Strike will be impossible. To that end, the US needs to pursue key technologies and protocols to protect it from attack and natural disasters.

If all of these recommendations can be accomplished, the US has a hope of enabling global strike in 2035. If it fails to overcome the obstacles, it will potentially be vulnerable to a devastating attack without the capability to respond, or deter others in the future.

Bibliography

Adhikari, Richard. "Massive Chinese Net Reroute Exposes Web's Achilles' Heel", *Technewsworld*, 17 Nov 2010. http://www.technewsworld.com/story/71258.html (accessed 20 Jan 2012).

Bell, D.E. and LaPadula, L. "Secure Computer Systems", ESD-TR-73-278, MITRE Corporation, vI and II, November 1973, and vIII, April 1974.

Blais, Andrew and Mertz, David. "An Introduction to Neural Networks." IBM Developer Works, 1 Jul 2001. http://www.ibm.com/developerworks/library/l-neural/ (accessed on 20 Jan 2012).

Bowman, Erik C. "Multi-level secure distributed architecture for Space Situation Awareness", AFIT/IC4/ENG/05-01, Air Force Institute of Technology, 2005

Craven, Phil. "Google's PageRank Explained", WebWorkshop. http://www.webworkshop.net/pagerank.html (accessed 20 Jan 2012).

Garreau, Joel. Radical Evolution. New York: Broadway Books, 2005.

Hayward, Matthew. "Motivation for Shor's algorithm", Illinois Mathematics and Science Academy. 26 Apr 2008. http://alumni.imsa.edu/~matth/quant/299/paper/node20.html (accessed 23 Jan 2012).

Howard, N., Ph.D, HDR. "Cyber Security Challenge: What is to follow the Social Network?" Presentation, International Conference of Cyber Security Asia, Singapore, 8 - 11 November, 2011.International Conference of Cyber Security Asia, Singapore (8 - 11 November, 2011).

Ferguson, John (principal investigator). "STTR Phase I: Ultrafast Response Transient Voltage Surge Suppressors." National Science Foundation Contract Award 80154, 2006.

Massachusetts Institute of Technology-Lincoln Labs. "Extended Space Sensors Architecture". Tech notes. www.ll.mit.edu/publications/technotes/TechNote_ESSA.pdf (accessed on 23 Jan 2012).

Rivest, R., Shamir, A., and Adelman, L. "A Method for Obtaining Digital Signatures and Public Key Cryptosystems." Communications of the ACM, February 1978.

Schraudolph, N. and Cummins, F. "Introduction to Neural Networks." Online Course. Istituto Dalle Molle di Studi sull'Intelligenza Artificiale Lugano, CH. http://www.idsia.ch/NNcourse/brain.html (accessed 12 Jan 2012)

University of Leicester. "How Does The Human Brain Work? New Ways To Better Understand How Our Brain Processes", Science Daily, 19 May 2009. http://www.sciencedaily.com/releases/2009/05/090519152559.htm (accessed 29 Jan 2012).

Washington State Department of Environmental Health. Office of Radiation Protection. "Electromagnetic Pulse (EMP)". http://www.doh.wa.gov/ehp/rp/factsheets/factsheets-htm/fs41elecpuls.htm (accessed on 23 Jan 2012).

Whitehead, N., Overton, M., LaBry, Z., Hamilton, F., and Leising, B. "Encrypting Chaos: Fractal Encryption", New Mexico Adventures in Supercomputing Challenge. Final Report, 2 Apr 2003.

Notes

[1] Dr. Newton Howard. "Cyber Security Challenge: What is to follow the Social Network?" (presentation, International Conference of Cyber Security Asia, Singapore, 8-11 November 2011). Also, CBS's TV Series "Person of Interest" is a fictional example of the potential for this type of application.

[2] Certification and Accreditation (C&A) is the primary challenge for DoD users, one private users do not have to contend with. The other major challenge is to add context so that it can be understood and assimilated by a PED system. These are both non-trivial tasks.

[3] Security restrictions/concerns such as "Need to Know" and the resultant Certification and Accreditation requirements are the primary challenges here.

[4] N. Scraudolph and F. Cummins. "Introduction to Neural Networks. Lecture 1: Computation in the Brain." Online course (Istituto Dalle Molle di Studi sull'Intelligenza Artificiale Lugano, CH), http://www.idsia.ch/NNcourse/brain.html (accessed 13 Jan 2012).

[5] Ibid

[6] Ibid

[7] Ibid

[8] University of Leicester. "How Does The Human Brain Work? New Ways To Better Understand How Our Brain Processes", *Science Daily*, 19 May 2009. http://www.sciencedaily.com/releases/2009/05/090519152559.htm (accessed 29 Jan 2012)

[9] Schraudolph and Cummins, "Computation in the Brain."

[10] Ibid.

[11] Ibid

[12] Ibid

[13] Ibid

[14] Ibid

[15] Note that the brain also sometimes recognizes patterns where a pattern does not really exist (e.g. conspiracy theories, cloud shapes, etc.). The system will also have to try to avoid drawing false conclusions.

[16] Note that the Internet does not truly emulate the brain because the Internet's domain name system (DNS) is limited to only 2 direct connections, at least when working with uniform resource locators (URL). In contrast, the brain has no defined upper limit on connections. The Internet Protocol (IP) routing scheme is generally more robust, but unlike the human brain, is vulnerable to purposeful manipulation. In April 2010, China "accidentally" rerouted all Internet traffic through China for 18 minutes by reloading the IP routing tables of key hubs on the internet. Richard Adhikari. "Massive Chinese Net Reroute Exposes Web's Achilles' Heel", *Technewsworld*, 17 Nov 2010 http://www.technewsworld.com/story/71258.html, (accessed 20 Jan 2012).

[17] The concept of an index to information is a key concept that will be critical to the operation of the ISR PED system—it cannot operate effectively without one.

[18] Phil Craven. "Google's PageRank explained", WebWorkshop. http://www.webworkshop.net/pagerank.html
(accessed 20 Jan 2012).

[19] For example, Archie Bunker or Edith Bunker.

[20] Andrew Blais and David Mertz. “An Introduction to Neural Networks.” IBM Developer Works, 1 Jul 2001. http://www.ibm.com/developerworks/library/l-neural/ (accessed 20 Jan 2012).

[21] These reasons include: they do not trust that other agencies will use the information appropriately; they fear other agencies will inadequately protect the information; or agencies simply want the power that comes with controlling the information. The first two reasons could result in compromising sensitive sources and methods for acquiring that information.

[22] If one could prove to data owners that the system that could adequately prevent compromise of data and sources, it might be easier to persuade them to share information. This is a primary reason for considering security as an absolute requirement of the PED system.

[23] This challenge makes it difficult to entice contractors to work as a system integrator. The profit is in developing new proprietary applications, not integrating existing systems. In order to integrate, the contractor must often sign non-disclosure and non-competition clauses to access application data from competitors, which cuts them out of lucrative application development work in the future.

[24] In the author’s acquisition experience, this process has taken upwards of 9 months per application. Entire programs have been cancelled or reassigned due to inability to actually deploy capabilities that were tested successfully on developmental systems. On SMC’s now-defunct Integrated Space Situation Awareness (ISSA) program, the program office was developing “glueware” to link over 40 applications. Although each glueware application could be developed in a matter of weeks, the program office calculated it would take approximately 33 years to sequentially deploy all of the capabilities based on approval timelines for the first applications that were developed.

[25] Although there are efforts to read and modify the minds of individuals through various techniques, it is only able to localize to regions of the brain, not individual neurons – yet. Joel Garreau, *Radical Evolution* (New York: Broadway Books, 2005.) 35-38.

[26] Humans do not share secrets with someone unless they recognize and trust that person. Computers need a way to recognize and trust users as well before allowing them access.

[27] Third party introductions are where a single recognizable person is trusted and introduces two strangers. The trusted third party then vouches for each person to the other. The key is both strangers recognize and trust the third party.

[28] Currently, computer authentication requires the use of digital signatures using public key cryptography. For specifics on how a PED might use cryptographic digital signatures, see Erik C. Bowman. “Multi-level Secure Distributed Architecture for Space Situation Awareness”, AFIT ENG/IC4/05-01, 2005, Appendix C.

[29] Ibid.

[30] The author acknowledges that by 2035 data formats will change, causing difficulties accessing and linking all data. But as with the rest of the system, the concept of translators or sidecars should be able to address legacy data formats until the data stored on older sources is updated into more current formats.

[31] Ronald Rivest, Adi Shamir, and Leonard Adelman. “A Method for Obtaining Digital Signatures and Public Key Cryptosystems.” Communications of the ACM, February 1978.

[32] This is a simplification. The public key and private keys are not exactly what are described, but this explanation is close enough to reality for illustrative purposes.

[33] Matthew Hayward. “Motivation for Shor’s algorithm”, Illinois Mathematics and Science Academy. 26 Apr 2008. http://alumni.imsa.edu/~matth/quant/299/paper/node20.html (accessed 23 Jan 2012). This (among other sources, including Wikipedia, which has a fairly accurate description) discusses how the use of Shor’s algorithm on a quantum computer can increase the speed of factorization exponentially compared to the most efficient classical factorization method, the general number field sieve.

[34] An excellent discussion of applicability of fractals and chaos to encryption can be found at Nick Whitehead, Michael Overton, Zach LaBry, Franklin Hamilton, and Brian Leising. “Encrypting Chaos: Fractal Encryption”, New Mexico Adventures in Supercomputing Challenge, Final Report, 2 Apr 2003. The potential applicability of fractal theory to encryption is based on the fact that quantum computers are best suited to solving problems classed in the category of “single instruction multiple data”. Factoring large numbers falls into this category. In contrast, fractals and chaotic systems are “multiple instruction, single data” problems.

[35] Washington State Department of Environmental Health. Office of Radiation Protection. “Electromagnetic Pulse (EMP)”. http://www.doh.wa.gov/ehp/rp/factsheets/factsheets-htm/fs41elecpuls.htm (accessed 23 Jan 2012).

[36] John Ferguson (principal investigator). “STTR Phase I: Ultrafast Response Transient Voltage Surge Suppressors.” National Science Foundation Contract Award 80154, 2006. http://www.sbir.gov/sbirsearch/detail/83355, (accessed 23 Jan 2012).

[37] One could also view this web as a tapestry. A single thread does not give one much awareness. Until the threads are woven into a tapestry, there will not be a complete picture.

[38] While neural nets could eventually be used, context, data and interface standards will initially need to be defined manually. Once defined, neural nets can be trained to automate the process in the future.

[39] The reason for developing the initial increment on SIPRNET is that it contains both unclassified and SECRET data on the same network. This enables one to fully develop, test, and implement the ability of the security model to segregate data by classification without running the risk that secret information could leak out onto the internet. Demonstration of no information leakage would pave the way for cross-network implementation.

[40] Massachusetts Institute of Technology-Lincoln Labs. “Extended Space Sensors Architecture” Tech notes, www.ll.mit.edu/publications/technotes/TechNote_ESSA.pdf (accessed 23 Jan 2012).

[41] Note that Google does this today, but as described in Section 2, does not leverage context. What is discussed here is more sophisticated, leveraging contextual information that is currently not embedded with context to make richer data linkages. But it employs the same basic tools as currently pioneered by Google and others.

[42] DE Bell and L. LaPadula. ”Secure Computer Systems”, ESD-TR-73-278, MITRE Corporation, vI and II (Nov 1973), vIII (apr 1974).

[43] Bowman. “Multi-level Secure Distributed Architecture for Space Situation Awareness”, J-1.

[44] Ibid, Appendix C.

[45] The idea of a cyber range, while beneficial for developing new capabilities, is not sufficient for integrating existing systems. The cyber range is a model of reality, not reality. Once the system is deployed into the operational environment, there is a high probability the

system won’t work. Examples this author has encountered are quirks in firewalls, proxy servers, device drivers, processing power, operating system version, level of patching, etc. In addition, the network continually changes—new nodes appear and disappear regularly.   Some level of maintenance will be required almost constantly to adapt to changes implemented by network operations. And of course, automated systems that dynamically modify links are in effect developing new capabilities on the operational network.